Working Paper No.125

# UNIT ROOTS IN SELECTED INDIAN MACROECONOMIC TIME SERIES



FAHIMUDDIN
B. BHASKARA RAO

GIRI INSTITUTE OF DEVELOPMENT STUDIES
Sector O, Aliganj Housing Scheme
LUCKNOW

# UNIT ROOTS IN SELECTED INDIAN MACROECONOMIC TIME SERIES*

FAHIMUDDIN

*Giri Institute of Development Studies, Lucknow*

B.BHASKARA RAO

*University of New South Wales, Sydney*

SUMMARY

A non-technical exposition of the unit roots and cointegration techniques is given. When unit root tests are applied to nine major macroeconomic time series of India, six frequently used variables are found to be non-stationary. A cointegrating equation between the price level, output and money and a short-run error correction model ($ECM$) are estimated. The cointegrating equation showed that while there is a strong long-run relationship between the price level and money, the relationship between the price level and output is weak. However the $ECM$ model showed that a strong short-run relationship between the rate of inflation and the rate of growth of output exists.

**JEL:** $E12$, $E13$.

**KEYWORDS:** Unit roots, cointegration, $ECM$, real output, money, price level.

* We are grateful to Professor V. K. Srivastava and all the participants in the Staff Seminar at the Giri Institute of Development Studies for their valuable suggestions and criticisms on this paper. Responsibility for errors remains with us only.

## 1. INTRODUCTION

Methodological revolutions in economics are not new. Economics has gone through the Keynesian, monetarist, neo-classical and new classical revolutions in the post war period. The major impact of each one of these revolutions is that they call for a fundamental change in our way of thinking about modelling economic phenomena. At the moment, both economics and econometrics are going through a new type of revolution viz. the unit roots and cointegration revolution. The new revolution also calls for a fundamental change in thinking about methods of estimation of economic relationships. However this new revolution does not seem to have made much impact on the applied econometric work in India partly due to the lack of adequate computer facilities and partly due to the neglect of the econometrics of time series models.[1]

In what follows, we shall

- provide an intuitive explanation of the significance of unit roots and cointegration,
- give an overview of the major steps involved in the applied econometric work with unit root variables,
- examine if there are unit roots in some selected Indian macroeconomic time series,
- use a simple method to estimate cointegrating equations to examine the long-run relationship between price level, real income and money supply and
- estimate a short-run dynamic relationship between these variables based on the error correction model (*ECM*).

It should be noted that our main objective in this paper is to illustrate the relevance and usefulness of these recent techniques and not to draw any definitive conclusions about the nature of the long and short-run relationships between the aforesaid variables. To achieve the latter objective satisfactorily more sophisticated techniques and data refinements are necessary; see Rao (1994) for some expository papers and references to the recent literature.

[1] However, a couple of recent works on the Indian economy have used some of these techniques; see Kalirajan and Shand (1994, pp.75-83) and Jadhav (1994, pp.238-239).

## 2. UNIT ROOTS AND COINTEGRATION

The standard classical methods of estimation, which we routinely use in the applied econometric work, are based on the assumption that the means, variances and autocovariances of the variables are well defined constants and *independent of time*. However, applications of the unit root tests have shown that these assumptions are not satisfied by a large number of macroeconomic time series. Variables whose means and variances change over time are known as non-stationary or unit root variables. Furthermore the unit root revolution also showed that standard classical estimation methods, such as the ordinary least squares ($OLS$), to estimate relationships with unit root variables give misleading inferences. This is known as the spurious regression problem and an intuitive explanation of its significance is as follows. If the means and variances of the unit root variables change over time all the computed statistics in a regression model, which use these means and variances, are also time dependent and fail to converge to their true values as the sample size increases. Furthermore conventional tests of hypothesis will be seriously biased towards rejecting the null hypothesis of no relationship between the dependent and independent variables. This is a serious problem if the null hypothesis is true.[2]

The economic implications of the unit roots literature are also equally profound. So far alternative macroeconomic paradigms have treated economic fluctuations as temporary deviations from a stable trend rate of growth of output and offered different explanations for these fluctuations. Therefore while there is bound to be some disagreement between these theories on the merits of alternative short-run stabilisation policies, such disagreements are less noticeable on policies to promote the trend growth.

Unit root tests, on the other hand, have shown that the assumption about the stability of the long-run trend rate of growth of output is untenable because aggregate output in many countries is found to be non-stationary. This finding casts a doubt on the usefulness of the ex-

---

[2] The spurious regression problem has other implications. Phillips (1986) developed a formal model for regressions between unit root variables to show, for example, that the $DW$ statistic converges towards zero. This is an important finding because low $DW$ statistics often indicate that the variables in a regression model are non-stationary. For a brief summary of Phillips' work see Ghosh (1994, pp. 534–535).

isting business cycle theories in which a distinction is made between the determinants of the trend rate of growth of output and cycles. Cointegration has also important implications. In this context it is convenient to view cointegration as a technique to estimate the equilibrium or long-run parameters in a relationship with unit root variables. Together unit roots and cointegration have important implications for the specification and estimations of dynamic economic models and call for a different strategy to model dynamic models. In the exiting strategies the methodological conflict between the equilibrium framework of the theory and the disequilibrium environment from which data are gathered is resolved by extending the equilibrium specifications to include disequilibrium adjustment mechanisms. The extended equation is then estimated from which estimates of the long-run or equilibrium parameters are derived by imposing equilibrium conditions. These equilibrium parameters are then used to test the underlying theory.

Even though the traditional approach can be criticised for neglecting the problems caused by the presence of unit root variables, the main advantage of cointegration is that it can be used directly to test or falsify the underlying theory. Furthermore cointegration enables utilisation of the estimated long-run parameters in the estimation of the short-run disequilibrium relationships. There is, however, a tradeoff. Estimation of cointegrating regressions and joint estimation of both the long and short-run specifications are computationally demanding. But this difficulty should not mask the significance of the unit roots and cointegartion revolution.

## 3. AN OVERVIEW

There are broadly four major steps in applying unit root and cointegration techniques.[3] First, unit root tests are applied to determine if the variables in a regression are stationary or nonstationary. Second, cointegrating regressions are estimated if the variables satisfy certain conditions. These cointegration regressions are the long-run or equi-

[3] Unfortunately there is no satisfactory and comprehensive discussion of these four steps in the existing econometrics textbooks which can be easily understood by an average applied economist. However, the interested reader may refer to Rao (1994) for various expository papers. Data used in these expository papers are also made available for easy replication of the results. Maddala (1992) has also an introductory discussion of these topics.

librium relationships between these variables. Third, the short-run or the dynamic disequilibrium relationships are estimated utilising the estimates of the long-run parameters within the error correction framework. Finally, the robustness of the estimated dynamic disequilibrium relationships is determined by subjecting them to the standard diagnostic tests.

Sometimes it might be possible to show that the unit root null hypothesis can be rejected for a set of variables. Needless to say this would considerably simplify estimation because all the standard classical methods can be utilised in these circumstances. For this purpose the unit root null hypothesis can be confronted with a stronger alternative hypothesis; see Perron (1989) and Rao (1993a, b) for applications of the Perron tests.

The aforesaid overview is a highly simplified and condensed version of the major steps involved in applying unit root and cointegration techniques. In practice, however, the applied economist will encounter several problems. Firstly, there is more than one method of testing for the presence of unit roots and to estimate the parameters in the cointegrating regressions. Secondly, the critical values for each of these tests and estimation methods depend upon the number of variables and also the specification selected for an equation. Nonetheless it can be said that the augmented Dickey Fuller test (*ADF*) is widely used to test for the presence of unit roots in the variables. A more powerful alternative is the Phillips and Perron non-parametric test. The Engle and Granger (1987) two step method is the simplest method of estimation of the cointegrating regression and error-correction formulation. However, the Johansen maximum likelihood method is now widely used for this purpose.[4]

## 4. UNIT ROOTS IN INDIAN DATA

We have selected for examination and illustration a small number of frequently used macroeconomic variables in the applied econometric

[4] Needless to say applied work in this area is hard to conduct without adequate computer hardware and software facilities. A *386 PC* with a co-processor and a good hard disk is the minimum requirement. Several well known statistical packages such as *Microfit, PcGive, RATS, SAS, TSP* etc. can be used to test for unit roots, estimate cointegrating regressions and error correction specifications. We have used *Microfit,* for computations in this paper.

work on the Indian economy. These are: real gross national product ($Y$), implicit $GNP$ price deflator ($P$), real output in the agricultural and industrial sectors ($Y_A, Y_I$), real saving and investment expenditures ($S, I$), real exports ($E$), real imports ($IM$) and nominal money supply ($M$). The real magnitudes of the variables are obtained by deflating the nominal values with the implicit $GNP$ deflator. Further details of these data are given in the Appendix. The period covered ranges from 1950-51 to 1993-94 but it is not the same for all the variables due to some data limitations.

It has been pointed out earlier that there are three types of tests for testing for unit roots but the augmented Dickey-Fuller test ($ADF$) is popular. We shall use only the $ADF$ test in this paper.[5] We present now a *simplified procedure* to implement the $ADF$ test to determine if there is a unit root in the variable $X$. For this purpose first estimate the following equation with $OLS$

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 T + \sum_{1}^{m} \Delta X_{t-m} \tag{1}$$

where $X$ is the variable under investigation, $T$ is time and $\Delta X$ is the first difference of $X$. In practice an investigator should add as many lagged first differences of $X$, i.e. increase $m$, as necessary to eliminate serial correlation in the residuals of equation (1).

The test statistic for the presence of unit root in $X$ is the $t$-statistic of the coefficient of $X_{t-1}$ i.e. the $t$ ratio of $\alpha_1$. Note that the null hypothesis is that $X$ is a unit root variable i.e. $X$ is a non-stationary variable and the test statistic and its critical value are both negative. Therefore if the *absolute* value of the test statistic is less than the *absolute* value of the tabulated critical value, the null hypothesis cannot be rejected and we conclude that $X$ is non-stationary.

---

[5] The other two tests are the Dickey-Fuller test and the Phillips-Perron ($PP$) non-parametric tests. The first of these assumes that the disturbance terms in the test equations are serially uncorrelated. In contrast both the $ADF$ and $PP$ tests allow for the possibility of stationary autoregressive error terms.

We have selected the $ADF$ test because it is much simpler to implement with any $OLS$ regression software package. Furthermore we followed the simpler approach of Dickey, Jansen and Thornton (1994) and did not use more systematic procedures based on the Dickey and Fuller (1981) $\Phi_i$ tests; see Holden and Perman (1994) for a sequential procedure.

Unfortunately this test statistic does not have the usual $t$ or normal distribution and the critical values are given in the bottom part of Table 8.5.2. in Fuller (1976, p.373) for sample sizes of 20, 50, 100, 250, 500 etc. Since our sample size is close to 50 it is worth noting that the 5% and 10% critical values are $-3.50$ and $-3.18$ respectively. Thus if the estimated *absolute* value of the $t$ statistic of $\alpha_1$ in our $ADF$ test equation is more than the *absolute* critical values we reject the null hypothesis that variable $X$ is non stationary.

Results based on the augmented Dicky-Fuller (ADF) test for these 9 variables are given in Table 1. *All variables are transformed into their natural logarithms.* In column 4, $m$ is the number of lags used on the first differences of the variable to eliminate serial correlation in the residuals. The corresponding $\chi^2$ statistics, for the null that there is no $4^{th}$ order serial correlation in the residuals of the equation, given in column 5, are all insignificant at the 5% level. The $t$ ratios of the estimated coefficients of the lagged variables (our unit root test statistics) are given in column 6. Critical values for the null of unit roots for the 5% and 10% levels are shown in columns 7 and 8 respectively.

It can be seen from these results that while three variables viz. real output in the agricultural sector ($Y_A$), real saving ($S$) and real investment ($I$) are stationary, the other six variables viz. real output ($Y$), real output in the industrial sector ($Y_I$), real exports ($E$), real imports ($IM$), price level ($P$) and money supply ($M$) are all non-stationary. Therefore while $OLS$ regression equations with the levels of the any three of the stationary variables are meaningful, such equations with any of the six non stationary variables yield misleading conclusions.

As examples it can be said that $OLS$ equation say between $I$ and $S$ is meaningful but $OLS$ equations say between $S$ and $Y$ or between $P, M, Y$ yield misleading results. To illustrate we give below $OLS$ equations between $I$ and $S$ and $S$ and $Y$. The first equation examines if a change in saving leads to a change in investment and the second equation which is more familiar is the saving function.[6] Note that all these variables are in their natural logarithms and therefore the estimated coefficients

[6] The problem of causality i.e. whether $I$ depends on $S$ or $S$ depends on $Y$ is interesting and worth examining. But this is beyond the scope of this paper. We simply assume that $I$ depends on $S$. Interesting empirical work on this topic where the unit root problem is also examined can be found in Kalirajan and Shand (1994) and Mehra (1994). Both papers also give data for easy replication.

Table 1

AUGMENTED DICKY-FULLER TESTS FOR THE UNIT ROOT HYPOTHESIS: LEVELS OF VARIABLES

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 T + \sum_1^m \gamma_m \Delta X_{t-m}$$

| | *Var* | *Period* | $m$ | $\chi^2_{se}$ | $t_{\alpha 1}$ | $CV_{F5\%}$ | $CV_{F10\%}$ |
|---|---|---|---|---|---|---|---|
| 1. | $Y$ | *1953-93* | 1 | 6.564 | −1.679 | −3.50 | −3.18 |
| 2. | $Y_A$ | *1953-93* | 1 | 1.294 | −4.155* | −3.50 | −3.18 |
| 3. | $Y_I$ | *1953-93* | 1 | 6.985 | −2.462 | −3.50 | −3.18 |
| 4. | $S$ | *1953-93* | 1 | 1.744 | −4.697* | −3.50 | −3.18 |
| 5. | $I$ | *1955-93* | 3 | 8.563 | −3.774* | −3.50 | −3.18 |
| 6. | $E$ | *1955-93* | 3 | 9.163 | −1.044 | −3.50 | −3.18 |
| 7. | $IM$ | *1953-93* | 1 | 2.096 | −2.014 | −3.50 | −3.18 |
| 8. | $P$ | *1956-93* | 4 | 5.526 | −2.676 | −3.50 | −3.18 |
| 9. | $M$ | *1953-93* | 1 | 9.135 | −1.353 | −3.50 | −3.18 |

Notes:

All variables are transformed into their natural logarithms.

* indicates significance at the 5% level.

The $\chi^2_{se}$ statistic, for testing $4^{th}$ order serial correlation in the residuals, has 4 degrees of freedom. The 95% critical value is 9.488.

$CV_{F5\%}$ and $CV_{F10\%}$ are the 95% and 90% critical values from the bottom section of Table 8.5.2 of Fuller (1976, p.373).

are elasticities.

$$I_t = 0.098 + 0.350\ S_t + 0.649 I_{t-1} \qquad (2)$$
$$(1.10) \quad (5.70)_* \quad (10.04)_*$$

$$R^2 = 0.999,\ \rho_1 = 0.563\ (3.77)_*,\ S.E.\ 0.039.$$

$$S_t = -9.250 + 1.053\ Y_t + 0.703 S_{t-1} \qquad (3)$$
$$(3.79)_* \quad (3.79)_* \quad (8.50)_*$$

$$R^2 = 0.997,\ h = 0.368,\ S.E.\ 0.090.$$

$t$-ratios are given in the parentheses. * indicates significance at the 5% level. The estimates of both equations seem satisfactory. Equation (2) with stationary variables has been estimated with the Cochrane-Orcutt method to eliminate first order serial correlation in its residuals.

The estimate of the serial correlation coefficient ($\hat{\rho}_1$) is significant. This equation implies that, at the means of the variables, a *Rs*.1.00 increase in saving leads to about *Rs*.0.95 increase in investment in the long-run.[7] Although estimates of the saving function in equation (3) seem satisfactory, the presence of $Y$, a unit root variable, gives raise to implausible estimates for the long-run marginal propensity to save of about 2.564; see footnote 7.

This does not, however, mean that there is no relationship between $S$ and $Y$ but this relationship cannot be captured with simple *OLS* regression methods. To estimate the long and short-run relationship between such variables one should use cointegration techniques.

## 5. COINTEGRATION

In what follows we shall use for illustrative purpose a simple procedure to obtain the cointegrating relationships between some unit root variables and then estimate the long and short-run relationships between these variables. For this purpose we select $M, Y$ and $P$ which are a part of the simple quantity theory of money.

We shall use a simple two step procedure developed by Engle and Granger (1987) to estimate the cointegrating and *ECM* equations between $P, M$ and $Y$; see Maddala (1992, pp.590-592) for a good introduction. In this context it is useful to view these cointegrating equations as the long-run relationships between the levels of these variables. As pointed out earlier the cointegration equation has no casual interpretation and one should depend on economic theory or the causality tests to determine what causes what.

Estimation of the cointegrating equations between a set of variables is possible only if all the variables in that equation are integrated to the same order. Furthermore, in principle, the number of cointegrating equations is one less than the number of variables in a relationship. A variable $X_t$ is said to be integrated to the order zero and denoted as $I(0)$ if it is a stationary variable i.e. there are no unit roots in that variable. Thus $I, S$ and $Y_{AG}$ are all $I(0)$ variables but the other six variables of Table 1 are not $I(0)$. The variable $X_t$ is said to be integrated to order one and denoted as $I(1)$ if $X_t$ is non-stationary but $\Delta X_t$ is stationary.

[7] The long-run elasticity of $I$ with respect to $S$ is $0.994 = 0.349/(1 - 0.649)$. The ratio of the mean of $I$ to $S$ is 0.956. Corresponding values for equation (3) are 3.550 and 0.722. These values imply the results in the text.

Similarly it will be an $I(2)$ variable if $\Delta X_t$ is non-stationary but $\Delta^2 X_t$ is stationary. It is easy to guess now that the order of a variable can be simply determined by applying, for example, the *ADF* test (or some other unit root test) to the first differences, second differences etc. of that variable in a sequence.

To determine the order of integration of the six unit root variables in Table 1, we apply the *ADF* test to their first differences. The results are given in Table 2 below.

**Table 2**

AUGMENTED DICKY-FULLER TESTS FOR THE UNIT ROOT HYPOTHESIS: FIRST DIFFERENCES OF VARIABLES

$$\Delta^2 X_t = \beta_0 + \beta_1 \Delta X_{t-1} + \beta_2 T + \textstyle\sum_1^j \theta_j \Delta^2 X_{t-j}$$

| | *Var* | *Period* | *j* | $\chi^2_{sc}$ | $t_{\beta 1}$ | $CV_{F5\%}$ | $CV_{F10\%}$ |
|---|---|---|---|---|---|---|---|
| 1. | $\Delta Y$ | *1954-93* | 1 | 6.790 | $-6.114_*$ | $-3.50$ | $-3.18$ |
| 2.† | $\Delta Y$ | *1955-93* | 2 | 3.774 | $-4.144_*$ | $-2.93$ | $-2.60$ |
| 3. | $\Delta Y_l$ | *1953-93* | 0 | 6.526 | $-4.825_*$ | $-3.50$ | $-3.18$ |
| 4. | $\Delta E$ | *1952-93* | 0 | 2.800 | $-5.576_*$ | $-3.50$ | $-3.18$ |
| 5. | $\Delta IM$ | *1952-93* | 0 | 8.478 | $-5.894_*$ | $-3.50$ | $-3.18$ |
| 6. | $\Delta P$ | *1956-93* | 3 | 2.773 | $-6.288_*$ | $-3.50$ | $-3.18$ |
| 7. | $\Delta M$ | *1954-93* | 1 | 8.165 | $-5.003_*$ | $-3.50$ | $-3.18$ |

Notes:

See notes for Table 1.

† Estimated without the time trend. Critical values for the test statistics of this type of equation are given in the middle section of Table 8.5.2 of Fuller (1976, p.373).

It can be seen from these results that all the six variables are found to be stationary in their first differences. Note that equation (2) in Table 2 is a variant of equation (1) where the trend variable $T$ is omitted. There is no hard and fast rule on whether the trend variable should or should not be omitted from the *ADF* equation.[8] Omission of the

[8] In Kalirajan and Shand (1994) the trend variable is omitted in the *ADF* tests on the first differences of the variables. However, there seems to be minor typographical error in the specification of the *ADF* equation (4.16). Instead of $\Delta x_{t-1}$, $x_{t-1}$ appears on the right hand side.

trend variable in the other equations does not change our conclusion and these results are not reported to conserve space. We may now conclude that all the six variables in Table 2 are $I(1)$.

For illustrative purpose we shall now determine if $Y, M$ and $P$ are cointegrated. Note that these three variables are all of the same order i.e. $I(1)$ and at the most there could be two cointegrating equations. To determine the number of cointegrating equations we should estimate with $OLS$ the following three linear equations because our simple procedure is sensitive with respect to the variable selected as the independent variable.

$$M_t = \beta_0 + \beta_1 Y_t + \beta_2 P_t \tag{4}$$

$$Y_t = \alpha_0 + \alpha_1 M_t + \alpha_2 P_t \tag{5}$$

$$P_t = \gamma_0 + \gamma_1 Y_t + \gamma_2 M_t \tag{6}$$

Note that these equations should not have lagged variables, first differences etc. and thus ignore the dynamics. There are two popular tests to determine if the estimated equation is a cointegrating equation. The first one is known as the cointegrating regression Durbin-Watson ($CRDW$) statistic and was developed by Saragan and Bhargava (1983). The null hypothesis is that there is no cointegration between the variables and the test statistic is simply the $D.W.$ statistic of the above $OLS$ equations. Critical values are given in Engle and Yoo (1987) and reproduced in Maddala (1992, p.607). If the computed $D.W.$ statistic is above the critical value, we reject the null hypothesis and conclude that the variables are cointegrated. The second test is the $ADF$ test applied to the residuals of the above $OLS$ equations. This test determines whether or not these residuals are $I(0)$. If the residuals are in fact $I(0)$ then the null hypothesis of no cointegration is rejected. The critical values are also given in Engle and Yoo (1987) and Maddala (1992).

In Table 3 we present the relevant coefficients and test statistics. The estimated coefficients are normalised on $M$ i.e. its coefficient is set to unity. It can be seen from these results that the $CRDW$ statistics indicate that there is no cointegration between these variables but the $ADF$ statistics show that there are in fact three cointegrating equations which is implausible. This calls for a more rigourous method, such as the Johansen maximum likelihood method, to test for cointegration between these variables but this is outside the scope of the present paper; see Dickey, Jansen and Thornton (1994) and Holden and Perman (1994) for an exposition of this method. In applied work, however, the

$ADF$ statistics is given more importance. On the basis of these results, since three cointegrating equations are implausible, we may say that perhaps the cointegration equation when $Y$ is the dependent variable i.e. the specification in equation (5) is least satisfactory because of its low $CRDW$ statistic and the closeness of its $ADF$ statistic to its critical value. Note that the specifications in equations (4) and (6) imply that they can be interpreted as the demand for nominal money and a simple quantity theory based price equation respectively. It is unusual to estimate the demand for money function in a nominal form. Therefore, in what follows, we shall use the quantity theory based price equation for further analysis because this relationship has also received attention by the Reserve Bank economists; see Rangarajan and Arif (1990) and Jadhav (1994).

**Table 3**

TESTS FOR COINTEGRATION

| | *Dep.Var.* | $Y$ | $P$ | $CRDW$ | $CV_{5\%}$ | $ADF$ | $CV_{5\%}$ |
|---|---|---|---|---|---|---|---|
| 1. | $M$ | 0.912 | 1.044 | 0.881 | 0.99 | $-4.037_*$ | $-3.75$ |
| 2. | $Y$ | 2.064 | 0.415 | 0.557 | 0.99 | $-3.995_*$ | $-3.75$ |
| 3. | $P$ | 0.374 | 1.370 | 0.807 | 0.99 | $-4.810_*$ | $-3.75$ |

Notes:

* implies significance at the 5% level.

$CV$ stands for tabulated critical values from Maddala (1992, p.607).

Normalised coefficients from the cointegrating regressions are reported columns 3 and 4.

## 6. ERROR CORRECTION MODEL

The $ECM$ formulation of an equation yields both the short and long-run specifications for the variable of interest. Its general form for our price equation is

$$\Delta P_t = \beta_0 + \text{Lagged values of } (\Delta Y, \Delta M, \Delta P) + \gamma_1(P_{t-1} - \gamma_2 Y_{t-1} - \gamma_3 M_{t-1}). \tag{7}$$

One retains as many lagged first differences in the above equation as are necessary to improve its diagnostic statistics. Thus its final form

is data based and needs careful judgement about the significance of various retained variables. Note that the parameters $\gamma_2$ and $\gamma_3$ are in fact the coefficients from the cointegrating equation given in (6) and its intercept $\gamma_0$ is subsumed into the new intercept $\beta_0$.

The above equation can be estimated with several alternative methods. However the Johansen maximum likelihood method in which equations (4) to (6) are seen as a system and all the equilibrium parameters are estimated simultaneously is the most efficient method. But, as noted earlier, it is computationally demanding and falls outside the scope of this paper. We shall use, therefore, the simpler Engle and Granger (1987) two-step procedure.[9]

In the Engle and Granger two-step procedure, the cointegrating regression (6) is first estimated with *OLS*. If its residuals are denoted as $\hat{Z}_t$, equation (7) can be expressed as:

$$\Delta P_t = \beta_0 + \text{lagged values of } (\Delta Y, \Delta M, \Delta P) + \gamma_1(\hat{Z}_{t-1}). \qquad (7')$$

This equation can be estimated in the second stage with *OLS* and its final form can be determined by retaining as many first differences of the variables on the right hand side as are necessary to improve the diagnostic statistics. Equations (8) and (9) below give the estimates of these two equations.

$$\begin{aligned} P_t = {} & 0.423 - 0.273\ Y_t + 0.730\ M_t \\ & (0.29) \quad (1.52) \quad\quad (11.31)_* \\ R^2 = {} & 0.994 \quad D.W. = 0.807 \end{aligned} \qquad (8)$$

$$\begin{aligned} \Delta P_t = {} & 0.032 - 0.307\Delta Y_{t-1} - 0.381\Delta Y_{t-2} + 0.391\Delta M_{t-1} \\ & (1.82)_{**} \quad (1.51) \qquad (1.74)_{**} \qquad (2.98)_* \\ & + 0.331\Delta P_{t-2} - 0.351\hat{Z}_{t-1} \\ & \quad (2.46)_* \qquad\quad (2.97)_* \end{aligned} \qquad (9)$$

$$R^2 = 0.443, \quad D.W. = 1.734, \quad S.E.E = 0.038$$

$$\chi^2_{sc} = 0.675, \quad \chi^2_{ff} = 1.472, \quad \chi^2_n = 0.910, \quad \chi^2_{hs} = 0.015$$

[9] This procedure is discussed in Maddala (1992, pp.262-264) together with similar procedures by others.

Significance at the 5% and 10% levels is indicated by * and **. In the cointegrating equation (8) while the coefficient of money is highly significant, the coefficient of real output is insignificant even at the 10% level. However it is significant at about 13% and has the expected sign. The $\chi^2$ statistics for equation (9) are for the null hypotheses that first order serial correlation ($\chi^2_{sc}$), functional form misspecification ($\chi^2_{ff}$), non normality of the residuals ($\chi^2_n$) and heteroscedasticity ($\chi^2_{hs}$) are absent and all these statistics are insignificant at the 5% level. The coefficient of the error correction term ($\gamma_1$) has the expected negative sign and significant. Its negative sign implies that any disequilibrium between the price level and its determinants is self correcting. The other coeffcients also have expected signs and significant. Thus our *ECM* equation can be said to be quite robust and satisfactory.

On the basis of these results it can be said that the long-run relationship between the price level and money is very strong. A ten percent increase in money leads to slightly more than seven percent increase in the price level in the long run. However the long-run relationship between the price level and real output is found to be week. This might be due to three reasons. Firstly, we have implicitly assumed that the velocity of circulation of money is constant and ignored its determinants e.g. the rate of interest. This may not be a valid assumption. Secondly, there seems to be some multicollinearty between *Y* and *M* and therefore it is hard to attach much meaning to the estimates of the individual coefficients. Thirdly, the Engle and Granger procedure is not very efficient and utilisation of more efficient methods of estimation, such as the Johanson maximum likelihood procedure, might throw a different light on the significance of real output. However, it should be noted from the *ECM* equation (9) that there is a strong short-run relationship between the rate of inflation and the rate of growth of output. The dynamics of the short run *ECM* equation is complicated. But it can be said that if the economy has already adjusted to equilibrium, a one percent increase in the rate of growth of income causes nearly a one percent decrease in the rate of inflation within a two year period.

## 7. CONCLUSIONS

In this paper we have first shown that there are unit roots in some important macroeconomic variable in India and then explained how regression equations with such variables should be estimated utilising

the unit root tests, cointegration techniques and *ECM* specifications. Our illustrative example examined the relationship between the price level, output and money supply and found that while the long-run relationship between the price level and money is strong, the corresponding relationship between price level and real output could not be adequately captured with our simple techniques. However our *ECM* equation found that there is a strong short-run relationship between the rate of inflation, and the rates of growth of output and money supply.

Needless to say there are several limitations in our paper some of which have been already noted. Firstly, there is a need to test for the presence of unit roots and estimate the long and short-run relationships between several other macroeconomic variables using various other alternative unit root tests. Secondly, although the Engle and Granger method of estimation yielded plausible preliminary estimates the usefulness of more efficient methods of estimation, such as the non-linear methods of estimation of the cointegrating and *ECM* equations with cross equation restrictions and the Johansen maximum likelihood method, should be examined. Thirdly, we did not pay much attention to refine the basic data or extend the estimation period. Therefore extensions to larger sample periods and examination of the data would be valuable because tests of unit roots and cointegration methods are large sample tests. Fourthly, in our illustrative example we have ignored the determinants of the velocity of circulation. Fifthly, the problem of multicollinearty between the explanatory variables in the cointegrating equation is not adequately addressed in the exiting literature. This is important because the variables in a cointegrating regressions are in their levels and there is bound to be some multicollinearty between such variables. Finally it is also worth investigating if the unit root null hypothesis can be rejected if it is tested against the Perron (1989) alternative hypothesis with structural breaks.

It is hoped that our preliminary analysis of the unit root problem with Indian data will attract more systematic and deeper analyses by other investigators.

## DATA APPENDIX

$Y$ = Real G.N.P. at factor cost in 1980-81 prices).
$P$ = G.N.P. deflator, 1980-81=100.
$Y_A$ = Real output in agricultural, forestry and logging, fishing, mining and quarrying, in 1980-81 prices.
$Y_I$ = Real output in manufacturing, construction, electricity, gas and water supply, in 1980-81 prices.
$S$ = Saving in the household, private corporate and public sectors. Nominal values are deflated by $P$.
$I$ = Gross domestic capital formation in the public and private sectors. Nominal values have been deflated by $P$.
$E$ = Exports including re-exports . Nominal values have been deflated by $P$.
$IM$ = Imports . Nominal values have been deflated by $P$.
$M$ = M1 definition of money supply.

All variables are measured in crores of rupees and transformed into their natural logarithms.

**Sources:**

For all variables except $M$ *Economic Survey* (New Delhi, 1993-94: Government of India). For money supply *Currency and Finance Report* (Bombay: Reserve Bank of India) various annual issues.

## REFERENCES

Dickey, D.A., Jansen, D.W. and Thornton, D.L. (1994) 'A Primer on Cointegration with an Application to Money and Income', in Rao, B. B. (ed.) *Cointegration for the Applied Economist* (Basingstoke: Macmillan), pp.9-46.

Duck, N.W. (1992) 'U.K. Evidence on Breaking Trend Functions', *Oxford Economic Papers*, 44, 426-439.

Engle, R.F. and Granger, C.W.J. (1987) 'Co-integration and Error Correction: Representation, Estimation and Testing', *Econometrica*, **55**, 251-276.

Engle, R.F. and Yoo, B.S. (1987) 'Forecasting and Testing in Cointegrated Systems', *Journal of Econometrics*, **35**, 143-159.

Fuller, W.A. (1976) *Introduction to Statistical Time Series* (New York: Wiley).

Ghosh, S. K. (1994) *Econometrics: Theory and Applications*, (New Delhi, 1994: Prentice-Hall of India.)

Holden, D. and Perman, R. (1994) 'Unit Roots and Cointegration for the Economist', in Rao, B. B. (ed.) *Cointegration for the Applied Economist* (Basingstoke: Macmillan), pp.47-112.

Jadhav, N. (1994) *Monetary Economics for India* (Delhi: Macmillan India).

Kalirajan, K.P. and Shand, R.T. (1994) *Economics in Disequilibrium* (Delhi: Macmillan India).

Maddala, G. (1992) *Introduction to Econometrics* (New York: Macmillan) 2nd ed.

Mehra, Y.P. (1989) 'Wage Growth and the Inflation Process: An Empirical Approach', in Rao, B. B. (ed.) *Cointegration for the Applied Economist* (Basingstoke: Macmillan), pp.147-159.

Microfit 3.0. (1991) *Users Manual* (Oxford: Oxford University Press).

Perron, P. (1989) "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis', *Econometrica*, **57**, 1361-1401.

——— (1994) "Trend, Unit Root and Structural Change in Macroeco-

nomic Time Series', in Rao, B. B. (ed.) *Cointegration for the Applied Economist* (Basingstoke: Macmillan), pp.113-146.

Phillips, P.C.B. (1986) 'Understanding Spurious Regressions in Econometrics', *Journal of Econometrics*, **33**, 311-340.

Ranjarajan, C. and Arif, R.R. (1990) 'Money, Output and Prices: A Macro Econometric Model', *Economic and Political Weekly,* April 21, 1990, pp.837-852.

Rao, B.B. (1993a) 'Unit Root Hypothesis, New Classical and Keynesian Models', *Economics Letters*, **41**, 47-52.

(1993b) 'The Nature of Transactions in the U.S. Aggregate Goods Market', *Economics Letters*, **41**, 385-390.

- (1994) *Cointegration for the Applied Economist* (Basingstoke: Macmillan).